**Alhasan A, Izuogu OG, Al-Balool HH, Steyn JS, Evans A, Colzani M, Ghevaert C, Mountford JC, Marenah L, Elliott DJ, Santibanez-Koref M, Jackson MS.**

**Circular RNA enrichment in platelets is a signature of transcriptome degradation.**

***Blood* 2016, 127(9), E1-E11.**

# Circular RNA enrichment in platelets is a signature of transcriptome degradation

Abd A. Alhasan[1#], Osagie G. Izuogu[1#], Haya H. Al-Balool[2], Jannetta S. Steyn[1,3], Amanda Evans[4], Maria Colzani[4], Cedric Ghevaert[4], Joanne C. Mountford[5,6], Lamin Marenah[6], David J. Elliott[1], Mauro Santibanez-Koref[1] and Michael S. Jackson[1]*

[1]Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK

[2]Kuwait Medical Genetics Centre, Al-Sabah Medical Area, Kuwait City, Kuwait

[3]Bioinformatics Support Unit, Faculty of Medical Sciences, Newcastle University

[4]NHS Blood and Transplant, Department of Haematology, University of Cambridge, Cambridge, UK

[5]Institute of Cardiovascular and Medical Sciences, University of Glasgow. Glasgow, UK

[6]Tissue and Cellular Therapies RD&I, Scottish National Blood Transfusion Service, Glasgow, UK

[#]These authors contributed equally to this work

*To whom correspondence should be addressed

**Key Points:**

- Circular RNAs are hugely enriched in platelets compared to nucleated cell types
- Lack of enrichment in megakaryocte progenitors implicates degradation of platelet linear RNAs

**ABSTRACT**

In platelets, splicing and translation occur in the absence of a nucleus. However, the integrity and stability of mRNAs derived from megakaryocyte progenitor cells remain poorly quantified on a transcriptome-wide level. As circular RNAs (circRNAs) are resistant to degradation by exonucleases, their abundance relative to linear RNAs can be used as a surrogate marker for mRNA stability in the absence of transcription. Here we show that circRNAs are enriched in human platelets 17-188 fold relative to nucleated tissues, and 14-26 fold relative to samples digested with RNAseR to selectively remove linear RNA. We compare RNAseq read depths inside and outside circRNAs to provide *in silico* evidence of transcript circularity, show that exons within circRNAs are enriched on average 12.7X in platelets relative to nucleated tissues, and identify 3162 genes significantly enriched for circRNAs, including some where all RNAseq reads appear to be derived from circular molecules. We also confirm that this is a feature of other anucleate cells through transcriptome sequencing of mature erythrocytes, demonstrate that circRNAs are not enriched in cultured megakaryocytes, and that linear RNAs decay more rapidly than circRNAs in platelet preparations. Collectively, these results suggest that circulating platelets have lost over 90% of their progenitor mRNAs, and that translation in platelets occurrs against the backdrop of a highly degraded transcriptome. Finally, we find that transcripts previously classified as products of reverse transcriptase template switching are both enriched in platelets and resistant to decay, countering the recent suggestion that up to 50% of rearranged RNAs are artefacts.

**INTRODUCTION**

Platelets are small, anucleate, circulating blood cells derived from large polyploid megakaryocyte progenitors [1]. They are vital for basic functions such as hemostasis, wound healing and angiogenesis [2]. Unlike anucleate erythrocytes which have a lifespan of ~4 months and have greatly reduced ribosome numbers [3], platelets are short lived (~8-11 days [4]), translationally competent [5], and some genes have been shown to undergo cytoplasmic splicing upon platelet activation [6, 7]. Consistent with these observations, recent RNAseq analyses have defined a complex transcriptome [8-11]. However, some transcripts believed to contribute to platelet function are expressed at very low levels in RNAseq data [11], and attempts to integrate transcriptome and proteome data have given conflicting results [12-15]. In the absence of nuclear transcription, RNAs within platelets are assumed to degrade over time. Although the abundance of individual mRNAs have been shown to fall in platelets stored at ambient temperatures [16], the stability and integrity of mRNAs within circulating platelets is poorly characterised at the transcriptome level. However, this is highly relevant to functional inferences drawn from RNAseq data [9, 11, 15, 17, 18].

Transcripts with rearranged exon order are now known to be very common in eukaryotic organisms, and are characterised at the sequence level by “back-splice” exon junctions inconsistent with the underlying genomic DNA [19-24]. Evidence for linear and polyadenylated structures has been presented [20, 25, 26], but it is now clear that most are both circular (circRNAs) and cytoplasmic [21-23]. Recent experiments using minigene expression constructs have established that intronic repeats flanking rearranged exons can promote circularisation, and that circRNAs primarily utilise canonical splice sites [27-30] and can be regulated by RNA binding proteins [29, 31]. Although two non-coding cytoplasmic circRNAs have been shown to act as miRNA sponges [22, 26], and circRNAs with retained introns (Exon-Intron or EIcircRNAs) can enhance transcription of their parental genes through interaction with U1 snRNP and RNA PolII [32], the vast majority of circRNAs consist of exons from protein encoding genes and have no known function [22, 33].

Circular RNAs are resistant to attack by exonucleases involved in the regulation of mRNA [34] and resistance to one, RNAseR, has been used to define them [23, 35]. Furthermore, blocking of transcription leads to an increase in circRNA levels relative to linear [23]. This differential stability raises the possibility that the abundance of circRNAs relative to linear forms could provide insight into the platelet transcriptome. Here, we analyse the circRNA population of platelets, and compare it to populations in nucleated cells, enucleated erythrocytes, and cells where linear RNA has been digested with RNAseR. We show that circRNAs are highly enriched in both platelets and erythrocytes relative to nucleated tissues, and confirm this experimentally using qPCR. We then analyse platelet

RNAseq read counts within exons which contribute to circRNAs and identify over 3000 genes where circRNAs are significantly enriched relative to nucleated cell types, including some genes where only circRNAs producing exons are represented within the RNAseq data. We also use qPCR to establish that circRNAs are not enriched in cultured magakaryocytes and that, over time, the levels of linear RNAs fall relative to circRNAs in platelet preparations. Our results suggest that over 90% of mRNAs have been lost from circulating platelets, and have potentially important implications for both platelet biology and our understanding of circRNAs.

## MATERIAL AND METHODS

*Datasets used*

*In silico* analyses (RNAseq, circRNA abundance and decay analysis) were performed on RNAseq data generated during the study and on 36 publicly available datasets. 24 of these were ribosomal RNA depleted total RNA fractions sequenced on the Illumina platform [11, 23, 27, 36, 37], 2 polyA+ selected samples sequenced on the Ilumina platform [11, 38], and 10 non-ribosomal RNA depleted total RNA sequenced on the SOLiD platform [15]. For details, see Supplementary Methods and Supplementary Table S1.

*Erythrocyte RNAseq*

Ribosome depleted RNAseq of total RNA from Red Blood cells was performed by AROS Applied Biotechnology (Aarhus, Denmark) using the TruSeq RiboZero Stranded mRNA LT kit (Illumina). The library was sequenced using a paired-end 100 base pair protocol on a Hi Seq 2000 (Illumina), all according to manufacturer's recommendations (see Supplementary Methods).

*circRNA identification*

Identification of back-splice exon junctions was performed using PTESFinder Pipeline v. 1 (http://sourceforge.net/projects/ptesfinder-v1/). This identifies all RNAseq reads which map in inverted order but the same orientation to the same RefSeq entry, generates models of these junctions, and re-maps all reads under analysis to the models. It then eliminates potential false positives through removal of reads which do not utilise known exon junctions, and reads which map equally well or better to the genomic or transcriptome reference (see Supplementary Methods and Supplementary Table S2 for details).

*Biological samples: origin and preparation*

RNAs from human solid tissues were obtained from Biochain (Amsbio, UK). Red Blood Cells were obtained from the Scottish National Blood Transfusion Service (SNBTS) in accordance with the terms of the standard donor consent form and information, and with approval of the SNBTS Sample Governance Committee. Cord blood was obtained after informed consent under a protocol approved by the National Research Ethics Service (Cambridgeshire 4 Research Ethics Committee ref. No. 07/MRES/44). Whole blood samples were collected from healthy volunteers with approval from Newcastle University's Faculty of Medical Sciences Ethics Committee. Platelets and platelet rich

plasma (PRP) were isolated according to a protocol which utilises a platelet inhibitory cocktail [39]. Isolation of peripheral blood mononuclear cells (PBMCs) and red blood cells (RBCs) was performed using Histopaque-1.077 density gradient (Sigma-Aldrich) according to manufacturer's instructions. Megakaryoctes were cultured and isolated according to the protocol of Tijssen et al. [40]. For further details see Supplementary Methods.

*RNA isolation and cDNA generation*

RNA was isolated using Trizol (Lifetechnologies) and treated with DNase I (Promega) according to manufacturer's instructions. RNA quantitation was performed using the NanoDrop ND-1000 spectrophotometer (Thermo Scientific) and quality was assessed using the Agilent 2100 Bioanalyser (Agilent Technologies). cDNA was synthesised using high capacity cDNA kits (Applied Biosystem), random hexamers, and Moloney murine leukaemia virus (MMLV) unless stated otherwise.

*Real time PCR*

Real time PCR was performed using Taqman mastermix (Lifetechnologies) as described previously [20]. Where linear and circular structures from the same gene were analysed, a common probe was used. Primers and assay efficiencies are given in Supplementary Table S1. Transcript expression was normalised using the ΔCT method relative to the geometric mean of 4 housekeeping genes (*GAPDH*, *PPIA*, *TUBB*, *GUSB*) analysed using TaqMan gene expression assays (Applied Biosystems). Circular transcripts were also normalised against the linear transcript from the same gene where appropriate. Reactions were performed in 10 µl volumes in 384 plates using QuantStudio 7 Flex (Applied Biosystems) with the following cycling parameters: 2 minutes at 50$^{o}$C, 10 minutes at 95 $^{o}$C, followed by 40 cycles of 15 seconds at 95 $^{o}$C and 60 $^{o}$C for 1 minute. The Wilcoxon rank sum test was used to compare levels of circRNA v linear RNA across multiple structures (analysis of paired $-\Delta$Ct values), and to compare linear v circRNA expression ratios between conditions or sample types (analysis of $-\Delta\Delta$Ct values across all structures).

## RESULTS

### Platelets are highly enriched for circRNAs

We first identified 3 publicly available RNAseq datasets from human platelets [9, 11, 15], one of which had a sequence read length long enough for the identification of back-splice junctions within circRNAs, and included samples processed to retain unpolyadenylated RNAs (Array express i.d. E-MTAB-1846; 100bp Illumina ribosome depleted total RNAseq from 3 individuals, and polyA+ RNAseq from 1 individual [11]). Analysis of the 3 total RNAseq samples using a back-splice junction discovery pipeline (PTESFinder see Methods) identified 33829 distinct structures consistent with circRNAs (Supplementary Figure 1 and Supplementary Table S3), approximately 4-5X more than reported for Fibroblasts [23] and ~15X more than reported for leukocytes and HEK293 cells [22] in libraries of comparable size. Furthermore, it identified 62%-64% of structures previously identified within RNAseq data from cell lines digested with RNAseR to selectively remove linear transcripts [23, 27], consistent with them being primarily derived from circRNAs. Despite the large number identified, the majority are not platelet specific, as 24632 (~73%) were also present in one or more of 15 ENCODE RNAseq datasets previously mined by Salzman et al. (2013) (Supplementary Figure S1).

Having identified an apparent abundance of circRNAs in platelets, we then used PTESFinder to identify circRNAs within other publicly available datasets generated using ribosome depleted total RNAseq and sequenced using the Illumina platform. These consisted of a panel of 12 human tissues [37], ENCODE cell lines [41], and cell lines treated with RNAseR to selectively remove linear RNAs [23, 27]. To enable comparison with a distinct anucleate cell type, we also generated >120 million 100bp paired-end Illumina RNAseq reads from total RNA extracted from leukocyte depleted red blood cells (RBCs, GEO record # GSE69192, see methods). These data, summarised in Figure 1 and presented in Supplementary Table S3, show that 17-188 times more circRNA (back-splice) reads are present in platelets than in human nucleated tissues (7579-11138 circRNA junction reads per million reads versus 59-447), and 14-26 times more circRNA reads than samples treated with RNAseR (427-463 junction reads per million reads). However, the numbers of circRNA structures and circRNA producing genes identified in anucleate cells are comparable to the numbers in RNAseR digested samples. In addition, the mean numbers of circRNA structures per circRNA producing gene are larger in anucleate and RNAseR digested samples than in others, with two of the three platelet samples averaging over 5 structures per gene, compared to 1-2 for nucleated tissues and cell lines.

We also analysed the platelet polyA+ sample from E-MTAB-1846 and identified 841 circRNA structures from 453 genes. However, their nucleotide composition was found to be significantly

enriched for A residues relative to structures present only in platelet total RNA (Chi squared = 12317 p<2.2e-16, Supplementary Table S4). As this suggests that these structures are non-polyadenylated circRNA transcripts rich in A nucleotides, they were not analysed further.

**More than 600 platelet genes generate over 20 distinct circRNAs.**

To illustrate the variation in circRNA distributions, the numbers of circRNA structures identified per gene are displayed for 4 samples in Figure 2A. Strikingly, both platelets and the fibroblasts digested with RNAseR (which increases circRNA levels by degrading liner RNAs) have large numbers of genes generating >15 distinct circRNA structures. For example, the platelet sample has >600 genes each with >20 different circRNAs. Furthermore, as shown in Table 1, individual structures in anucleate samples are supported by much larger back-splice junction read counts than in nucleated tissues ($p=2.7x10^{-4}$ anucleate v nucleated counts, Wicoxon rank sum test). All 3 platelet samples have over 1000 structures supported by over 100 reads, in sharp contrast to undigested nucleated samples. As an example of the diversity of structures and read support within platelets, the 49 inferred circRNA structures identified within the *XPO1* gene are shown in Figure 2B, with supporting read counts ranging from 1 (E8-E5 and E16-E7) to 4045 (E4-E2). It is noteworthy that all 21 of the internal exons of this gene are encompassed by one or more circRNA (illustrating that most exons within some genes can contribute to circRNAs), and that only 11 of these structures were identified previously in RNAseR digested H9 RNA [27].

**Real Time PCR confirmation of high circRNA transcript levels in anucleate cells.**

To experimentally confirm the enrichment of circRNAs in both anucleate cell types, we used qPCR to assess the relative levels of linear and circRNAs. We first assayed expression of the mononucleocyte expressed CD45 antigen (*PTPRC*), and performed cell counts to confirm that nucleated white blood cells were being efficiently removed during platelet isolation (Supplementary Figure S2 and Supplementary Methods). As a precaution against Reverse Transcriptase (RT) specific artefacts which have been reported [42] we also checked that all circRNAs being analysed could be amplified using templates generated with both AMV and MMLV RTs (Supplementary Figure S3). We then analysed the abundance of two known circRNA structures from *MAN1A2* (E4-E2 and E5-E2 [22]) relative to the canonical *MAN1A2* transcript, two structured from *PHC3* (a gene expressed in platelets for which assays were already available [20]), and an additional 7 abundant circRNAs together with their associated linear RNAs (see Supplementary Table S1). For each, paired assays with a common probe were used (Figure 3A).

Both of the *MAN1A2* circRNA junctions are much more abundant in platelets and RBCs than the linear mRNA (Figure 3B). However, they are less abundant than the linear transcripts in virtually all other nucleated tissues analysed. This is consistent with both the RNAseq data and previous analyses of nucleated tissues [20, 22]. The other 9 circRNAs are also all expressed at higher levels in platelets and RBCs than their corresponding linear structures (Figure 3C), with circRNAs from *SMARCA5*, *UBXN7* and *PNN* registering ~6-10 cycles before their corresponding linear RNAs, suggesting that they are ~50-1000 fold more abundant. In contrast, all are expressed in nucleated cells at an equivalent or lower level than their linear counterparts. Collectively, the difference in relative expression levels of linear and circRNAs between anucleate and nucleated tissues is highly significant ($p=2.7x10^{-12}$ Wilcoxon rank sum test).

Finally, the impact of RNAseR digestion upon expression was also assessed (Figure 3D). Digestion increases the expression of all back splice junctions relative to the canonical linear junctions in each cell type analysed ($p=7.8x10^{-3}$ Wilcoxon paired rank sum test, mock digest v RNAseR digest), although the *MAN1A2* E5-E2 back-splice shows only a marginal increase in the platelet sample (Figure 3D). This confirms that these structures include circRNAs. However, the impact of RNAseR digestion upon the relative levels of linear and circRNAs is more pronounced in peripheral blood mononuclear cells (PBMC) and in HEK293 cells than in either anucleate sample ($p<0.02$, Wilcoxon paired rank sum test for all pairwise nucleated v anucleated comparisons).

**For some genes, only reads from circRNA producing exons are detected in platelets**

The abundance of individual circRNAs relative to mRNAs from the same loci has previously been estimated *in silico* by comparing back-splice junction read counts to flanking canonical junction counts [20, 21, 43]. However, because most circRNA producing genes generate multiple overlapping structures (e.g. Figure 2B and Supplementary Table S3), the canonical junctions flanking one circRNA are often present within others. We therefore performed a transcriptome-wide analysis where we compared the average read count within exons which contribute to circRNAs, to the read count in exons which do not (see Supplementary Table S5). For each gene we calculated the RPKM across all exons internal to all known circRNA structures ($RPKM_I$), and the RPKM across exons external to all known circRNAs ($RPKM_E$). The former is a measure of both linear and circRNA expression, the latter linear only. Box and whisker plots showing $RPKM_I/RPKM_E$ ratios for all circRNA producing genes in all samples are presented in Figure 4A. The ratios from individual genes extend over 6-8 orders of magnitude for each sample, but what is most striking is that the median $RPKM_I/RPKM_E$ ratios of the anucleate samples are approximately 12 times higher than nucleated tissues, and approximately 5 times higher than samples digested with RNAseR (Figure 4A). The distribution of RNAseq reads

within these genes therefore reflects the pattern of circRNA structures inferred from exon-junction read counts alone (Figures 1 and 2), and provides further evidence that the back-splice junctions are predominantly derived from circRNAs.

To identify genes significantly enriched for circRNAs in platelets, for each gene we then estimated the contribution of circRNA producing exons to total transcription by calculating $RPKM_I/(RPKM_I + RPKM_E)$ ratios, and compared ratios in the platelet samples to ratios in the 12 nucleated tissues using the Wilcoxon rank test (p-value = 0.05, Benjamin and Hochberg false discovery rate = 0.01). Out of 8041 circRNA producing genes analysed, 3162 showed significantly enriched transcription from circRNA producing exons in platelets (Supplementary Table S5). When the ratios from all genes with an average RPKM >1 in platelets are plotted (Figure 4B), it can be seen that for the vast majority of enriched genes the contribution of circRNA exons to total transcription is >0.6 in nucleated tissues, but >0.8 in platelets. Furthermore, transcription from circRNA producing exons for 457 genes exceeds an average of 99% in the platelet samples, and for 15 genes it is 100% in all 3 samples (Supplementary Table S5) indicating that all reads detected are from circRNA producing exons. As this assay is not dependent on exon junction counts, which require long sequence reads for accurate detection, we were also able to apply it to one of the SOLiD platelet RNAseq datasets [15], and confirmed the presence of similar patterns of circRNA exon enrichment (e.g. Supplementary Figure 4, and Supplementary Data File 1).

**Reads from circRNA producing exons are enriched in platelets by up to 3000 fold**

To estimate the magnitude of circRNA exon enrichment for each gene in platelets relative to other tissues, we further analysed genes with an average $RPKM_I$ >1 by normalising their mean platelet $RPKM_I/RPKM_E$ ratio against the mean $RPKM_I/RPKM_E$ ratio from all 12 nucleated samples. The average enrichment was 12.7X for all genes, and 22X for the 3162 significantly enriched genes, with the highest enrichment being 3590X for *TMEM181* (Figure 4C and Supplementary Table S5). To illustrate these extreme levels of enrichment, RNAseq read counts and circRNA junction read counts for 4 enriched genes from a single platelet sample are shown in Figure 5. The vast majority of platelet RNAseq reads from *PNN* (enrichment 2189X) map to exons 7 and 8 which correspond to the most abundant circRNA (Figure 5A); the majority of reads from *SLC20A2* (enrichment 152X) map to the abundant circRNAs involving exons 3-5 (Figure 5B); all reads from *POMT1* (enrichment 127X) map to exons within its only circRNA (E4-E3 Figure 5C), while for *PHC3* (enrichment 240X), the reads map largely to the abundant E6-E5 and E7-E5 structures (Figure 5D).

Although all of our *in silico* analyses point to enrichment specifically in anucleate cells, no total RNAseq data has been generated from megakaryocytes, the direct progenitors of these cells. For all genes with assays available, we therefore also analysed the relative levels of linear and circRNAs in megakaryocytes cultured from cord blood. These were harvested after 10 days, when >70% of cells were positive for CD41 and CD42 (See Supplementary Methods). The results, presented in Figure 5 panels E and F, establish that megakaryocytes have very similar relative levels of linear and circRNA compared to other nucleated cell types (Fig 2B); all circRNAs are expressed at a much lower relative to linear RNAs than in platelets ($p=1.4x10^{-6}$ Wilcoxon paired rank sum test across all structures). This is consistent with circRNA enrichment occurring after platelet formation.

**Confirmation of differential decay of linear and circRNAs in platelet preparations**

Human mRNAs are subject to a wide variety of decay pathways [44], and exhibit extensive variation in decay rates [45-47]. To investigate decay in platelets, we first correlated the change in transcript abundance between megakaryocyte PolyA+ RNA and platelet PolyA+ RNA with published estimates of mRNA half-life [45]. This identified a weak but highly significant correlation, despite the difference in cell types between datasets analysed, with short half-life genes showing a more pronounced reduction in transcript abundance (r=0.17, p=3.35e-44, Figure 6A). This is consistent with a transcriptome-wide role for mRNA decay in platelet expression levels.

To experimentally investigate differential decay, we then used Real Time PCR to analyse changes in linear and circRNA abundance in platelet rich plasma (PRP) over time (see Methods). Consistent with previous analysis [16] initial experiments suggested that all transcript levels remained stable at 4$^{o}$C, but that degradation of mRNA could be detected when incubation temperature was elevated to 37$^{o}$C (data not shown). We therefore assayed the relative abundance of linear and circRNAs in PRP from 3 individuals, before and after incubation at 37$^{o}$C for 72 and 96 hours (Figure 6B). The levels of all 4 housekeeping genes were reduced by incubation, with raw CT values increasing by 4-6 cycles (Figure 6B top left panel). Reduction was also observed when genes which are highly abundant in platelets were analysed (Supplementary Figure S5). Normalisation of the linear transcripts from circRNA producing genes against the housekeeping pool (Figure 6B, top right panel) suggests that there are differences in decay rates between genes, with *UBXN7*, *PNN* and *MAN1A2* transcripts being enriched relative to control housekeeping RNAs. Strikingly, however, circRNAs analysed show significant enrichment (2-6 fold, Figure 6B, lower panel,) relative to linear RNAs after incubation ($p<2x10^{-3}$ for both day 0 v 72 hours and day 0 v 96 hours, Wilcoxon rank sum

test across all structures). This confirms that linear structures are decaying more rapidly than circRNAs in platelets.

**DISCUSSION**

Platelets are anucleate yet retain the machinery required for splicing and translation, and there is extensive evidence that specific genes are translated in activated platelets [5-7, 48-50]. As a result, recent platelet transcriptome analyses have generated considerable interest [8, 17, 18]. However, here we have shown that the platelet transcriptome is significantly enriched for circRNAs as assayed using back-splice junction counts, real-time PCR, and an RNAseq comparative read depth approach. Back-spice reads are enriched 17-188 fold relative to nucleated tissues, and over 3000 genes have a statistically significant enrichment of reads from circRNA producing exons. Furthermore, for some genes only RNAseq reads derived from circRNA producing exons are detected within platelet samples, highlighting the impact this enrichment will have upon RNAseq based estimates of expression levels. We have also established that circRNAs are not enriched in cultured megakaryocytes compared to other nucleated cell types.

While we cannot formally rule out the possibility that some circRNA enrichment occurs at the transcriptional level late in megakaryocyte maturation, collectively our results suggest that enrichment occurs specifically in platelets, after the loss of nuclear influence. The similarly high levels of enrichment observed in anucleate RBCs are consistent with this, while the observed reduction of linear RNAs relative to circRNAs in platelets incubated at 37$^{o}$C confirms that differential decay occurs. The overall level of enrichment of RNAseq reads from circRNA producing exons is consistent with platelet mRNA levels being at least 90% lower than within nucleated tissues (average enrichment 12.7X for genes with $RPKM_{I}$>1). Although this is a rough and indirect estimate, it is likely to be conservative as it assumes no degradation of circRNAs. Furthermore, the fraction of translatable RNAs is likely to be even lower, as our analysis has only considered RNA abundance, not structural integrity. This is consistent with the suggestion that platelet lifespan may be determined by translational competence [51].

Degradation has been assumed to be a feature of the platelet transcriptome based purely on their anucleate nature [51]. Mitochondrially encoded genes are among the most highly expressed [52], levels of individual mRNAs have been shown to fall when platelet preparations are incubated at ambient temperature [16], and there is some evidence that translation is reduced in aged platelets relative to the whole population [51]. Consideration of mRNA half lives alone would suggest that mRNA degradation is likely to be substantial, unless platelet specific mechanisms exist to counteract it. If we assume the average age of circulating platelets is 4 days (and that functional separation from the nucleus occurs at day 0), then mRNA half-lives [45] would predict gene specific reductions in expression ranging from 1.2X to 720X, with a mean reduction of 14X. Our estimate of 12.7X based on

RNAseq read counts is broadly consistent with this, despite the fact that $RPKM_I/(RPKM_I + RPKM_E)$ ratios will vary between genes due to a variety of factors which are unrelated to circularity, including differential isoform abundance, overlapping UTRs between neighbouring genes, and the reduction in RNAseq read coverage close to transcript termini.

Extensive mRNA degradation provides a simple explanation for conflicting reports concerning the correlation between the platelet transcriptome and proteome [12, 13, 15, 53], and for the presence of proteins in platelets without corresponding mRNAs [15, 50]. It can also explain why many genes which have been shown to be spliced and/or translated in platelets, such as Tissue Factor (*F3*) [7, 54], *IL1B* [55], *SLC23A2* [56], *FGA*, *FGB*, *FGG*, *ALB* [57], and *BCL3* [58], are present at very low levels or not at all in some RNAseq datasets [10, 11]. Collectively, our results indicate that translation in platelets occurs against the backdrop of a rapidly degrading transcriptome, that platelets will differ in their translational potential in an age dependent manner, and that RNAseq analyses targeted towards intact species which are capped [15, 18] or polysome bound will be required to analyse translational potential in these cells effectively.

CircRNAs have been identified from RNAseq data by the back-splice exon junctions they contain, and such junctions have been taken *prima facie* as evidence of circularity [24, 33, 59, 60]. Our exon-partitioned analysis of read-depth provides additional *in silico* evidence that the structures which contain these junctions are primarily circular, as trans-splicing of independent linear pre-mRNA [20, 42] would affect exons both up and downstream of the back-spliced exons. The higher number of circRNAs identified in platelets compared to RNAseR digested total RNA is, however, surprising as RNAseR selectively digests linear transcripts. This suggests that RNAseR removal of linear RNAs *in vitro* may be a relatively crude tool when compared to decay *in vivo*. This could be due both to limitations of enzyme specificity, and the lability of RNA post extraction. The significant enrichment of A nucleotides within back-splice transcripts present within platelet polyA+ RNA is also noteworthy as it means their presence within oligodT selected RNA fractions does not necessarily imply polyadenylation, as has been suggested previously [20, 42]. This may adversely affect the accuracy of RNAseq based mRNA expression estimates of the genes involved.

It has recently been suggested that up to 50% of RNAs with back splice exon junction are reverse transcriptase artefacts. This was extrapolated from an analysis of 13 back-splice RNAs where the only 6 RNAs which could be independently validated were those which could be amplified using both AMV and MMLV reverse transcriptases [42]. However, among the 7 RNAs which could not be validated, 4 were intragenic back-splice structures and all are present within platelets: The back-splice counts of 3 (*PHC3* E6-E5, *CNTLN* E5-E3, *RERE* E3-E3) are enriched relative to other tissues

(~1000X, ~100X, ~4X), and read counts in exons between each back-splice are also enriched (240X, ~30X, 500X) indicative of circRNAs. Furthermore, the *PHC3* E6-E5 structure degrades more slowly than linear RNA, is resistant to RNAseR digestion [23], and in our hands can be amplified from cDNA generated using either MMLV or AMV (Supplementary Figure S3). These diverse lines of evidence suggest that there may not be a straightforward relationship between RT-specificity and artefacts as previously suggested [42].

Finally, cellular levels of circRNAs are known to be low in proliferating and neoplastic human cells [43, 60] but to increase during differentiation of ES cells [61] and during Drosophila development [62], suggesting that they may be a biomarker for ageing. Among nucleated tissues they are highest in both human and Drosophila neural tissues [43, 62] and are particularly enriched in synapses [43]. They are also enriched in exosomes [63]. These observations, together with the presence of miRNA binding sites within circRNAs [62], and the dynamic expression patterns of some [43], implies that they may be of functional importance. CircRNA levels can be modulated by intronic homology [23, 27, 28], and splicing factors such as Quaking [31] and Muscleblind [29]. However, our qPCR analysis of transcripts in MKs suggests that the extreme enrichment in platelets reported here is not due to modulation at the pre-mRNA level within progenitor cells. With no transcription in platelets it can, therefore, only be accounted for by differential degradation/decay. The differential degradation and very high levels of enrichment reported here suggests that, even if the relative levels of linear and circRNAs from each gene remain constant at the level of transcription, subtle changes in expression level, rates of mRNA export, changes in the cell cycle [60], changes in nuclear and cytoplasmic volumes, and distance from the nucleus could all alter relative abundance over time. Many reported differences in linear and circRNA levels could, therefore, be due to their stability alone.

**ACKNOWLEDGEMENTS**

The financial support from the Leverhulme Trust (grant RPG-2012-795 to MSJ, MSK and DE), BBSRC (studentship BB/J014516/1 to OI), and the Wellcome Trust (grants WT089225MA and WT089225/Z/09/Z to DE) is gratefully acknowledged.

**AUTHOR CONTRIBUTIONS**

OI performed the bioinformatic analyses of RNAseq data, circRNA abundance, circRNA exon enrichment, and RNA decay. AAA performed all cell purification and isolation, RNA extraction, Real Time PCR analyses and platelet time course analyses. HHAI-B developed all real time PCR assays, JSS provided assistance with the analysis of PolyA+ transcripts, and JM and LM provided the erythrocytes used to generate RNAseq data. AE, MC and CG, isolated human haematopoietic progenitors from cord blood, followed by differentiation to megakaryocytes, and sorting to purity. MSJ, MSK and DE conceived the project; and MSJ drafted the manuscript with input from all authors.

**DISCLOSURE OF CONFLICTS OF INTEREST**

The authors declare that there are no conflicts of interests.

**REFERENCES**


1. Deutsch, V.R. and A. Tomer, *Megakaryocyte development and platelet production.* Br J Haematol, 2006. **134**(5): p. 453-66.
2. Garraud, O. and F. Cognasse, *Are Platelets Cells? And if Yes, are They Immune Cells?* Front Immunol, 2015. **6**: p. 70.
3. Migliaccio, A.R., *Erythroblast enucleation.* Haematologica, 2010. **95**(12): p. 1985-8.
4. Harker, L.A., et al., *Effects of megakaryocyte growth and development factor on platelet production, platelet life span, and platelet function in healthy human volunteers.* Blood, 2000. **95**(8): p. 2514-22.
5. Weyrich, A.S., et al., *Protein synthesis by platelets: historical and new perspectives.* J Thromb Haemost, 2009. **7**(2): p. 241-6.
6. Denis, M.M., et al., *Escaping the nuclear confines: signal-dependent pre-mRNA splicing in anucleate platelets.* Cell, 2005. **122**(3): p. 379-91.
7. Schwertz, H., et al., *Signal-dependent splicing of tissue factor pre-mRNA modulates the thrombogenicity of human platelets.* J Exp Med, 2006. **203**(11): p. 2433-40.
8. Freedman, J.E., *A platelet transcriptome revolution.* Blood, 2011. **118**(14): p. 3760-1.
9. Rowley, J.W., et al., *Genome-wide RNA-seq analysis of human and mouse platelet transcriptomes.* 2011. **118**: p. 101-111.
10. Bray, P.F., et al., *The complex transcriptional landscape of the anucleate human platelet.* BMC Genomics, 2013. **14**: p. 1.
11. Kissopoulou, A., et al., *Next generation sequencing analysis of human platelet PolyA+ mRNAs and rRNA-depleted total RNA.* PloS one, 2013. **8**: p. e81809.
12. Burkhart, J.M., et al., *The first comprehensive and quantitative analysis of human platelet protein composition allows the comparative analysis of structural and functional pathways.* Blood, 2012. **120**(15): p. e73-82.
13. Rowley, J.W. and A.S. Weyrich, *Coordinate expression of transcripts and proteins in platelets.* Blood, 2013. **121**(26): p. 5255-6.
14. Geiger, J., et al., *Response: platelet transcriptome and proteome--relation rather than correlation.* Blood, 2013. **121**(26): p. 5257-8.
15. Londin, E.R., et al., *The human platelet: strong transcriptome correlations among individuals associate weakly with the platelet proteome.* Biol Direct, 2014. **9**: p. 3.
16. Stiegler, G., et al., *P-selectin mRNA is maintained in platelet concentrates stored at 4 degrees C.* Transfusion, 2009. **49**(5): p. 921-7.
17. Rowley, J.W., H. Schwertz, and A.S. Weyrich, *Platelet mRNA: the meaning behind the message.* Curr Opin Hematol, 2012. **19**(5): p. 385-91.
18. Schubert, S., A.S. Weyrich, and J.W. Rowley, *A tour through the transcriptional landscape of platelets.* Blood, 2014. **124**(4): p. 493-502.
19. Dixon, R.J., et al., *A genome-wide survey demonstrates widespread non-linear mRNA in expressed sequences from multiple species.* Nucleic Acids Research, 2005. **33**: p. 5904-5913.
20. Al-Balool, H.H., et al., *Post-transcriptional exon shuffling events in humans can be evolutionarily conserved and abundant.* Genome Res, 2011. **21**(11): p. 1788-99.
21. Salzman, J., et al., *Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types.* PLoS One, 2012. **7**(2): p. e30733.
22. Memczak, S., et al., *Circular RNAs are a large class of animal RNAs with regulatory potency.* Nature, 2013. **495**: p. 333-8.
23. Jeck, W.R., et al., *Circular RNAs are abundant, conserved, and associated with ALU repeats.* RNA, 2013. **19**: p. 141-157.
24. Wang, P.L., et al., *Circular RNA Is Expressed across the Eukaryotic Tree of Life.* PloS one, 2014. **9**: p. e90859.

25. Caudevilla, C., et al., *Natural trans-splicing in carnitine octanoyltransferase pre-mRNAs in rat liver.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**: p. 12185-90.
26. Wu, C.-S., et al., *Integrative transcriptome sequencing identifies trans-splicing events with important roles in human embryonic stem cell pluripotency.* Genome research, 2013.
27. Zhang, X.-O., et al., *Complementary Sequence-Mediated Exon Circularization.* Cell, 2014. **159**: p. 134-147.
28. Liang, D. and J.E. Wilusz, *Short intronic repeat sequences facilitate circular RNA production.* Genes & development, 2014.
29. Ashwal-Fluss, R., et al., *circRNA Biogenesis Competes with Pre-mRNA Splicing.* Molecular Cell, 2014: p. 1-12.
30. Starke, S., et al., *Exon Circularization Requires Canonical Splice Signals.* Cell reports, 2014: p. 1-9.
31. Conn, S.J., et al., *The RNA Binding Protein Quaking Regulates Formation of circRNAs.* Cell, 2015. **160**(6): p. 1125-34.
32. Li, Z., et al., *Exon-intron circular RNAs regulate transcription in the nucleus.* Nat Struct Mol Biol, 2015.
33. Guo, J.U., et al., *Expanded identification and characterization of mammalian circular RNAs.* Genome Biology, 2014. **15**: p. 409.
34. Schoenberg, D.R. and L.E. Maquat, *Regulation of cytoplasmic mRNA decay.* Nature reviews. Genetics, 2012. **13**: p. 246-59.
35. Jeck, W.R. and N.E. Sharpless, *Detecting and characterizing circular RNAs.* Nature Biotechnology, 2014. **32**: p. 453-461.
36. Birney, E., et al., *Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project.* Nature, 2007. **447**: p. 799-816.
37. Nielsen, M.M., et al., *Identification of expressed and conserved human noncoding RNAs.* RNA, 2014. **20**(2): p. 236-51.
38. Nurnberg, S.T., et al., *A GWAS sequence variant for platelet volume marks an alternative DNM3 promoter in megakaryocytes near a MEIS1 binding site.* Blood, 2012. **120**(24): p. 4859-68.
39. Amisten, S., *A rapid and efficient platelet purification protocol for platelet gene expression studies.* Methods Mol Biol, 2012. **788**: p. 155-72.
40. Tijssen, M.R., et al., *Genome-wide analysis of simultaneous GATA1/2, RUNX1, FLI1, and SCL binding in megakaryocytes identifies hematopoietic regulators.* Dev Cell, 2011. **20**(5): p. 597-609.
41. Bernstein, B.E., et al., *An integrated encyclopedia of DNA elements in the human genome.* Nature, 2012. **489**: p. 57-74.
42. Yu, C.-Y., et al., *Is an observed non-co-linear RNA product spliced in trans, in cis or just in vitro?* Nucleic acids research, 2014. **42**: p. 9410-23.
43. Rybak-Wolf, A., et al., *Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed.* Mol Cell, 2015.
44. Schoenberg, D.R. and L.E. Maquat, *Regulation of cytoplasmic mRNA decay.* Nat Rev Genet, 2012. **13**(4): p. 246-59.
45. Friedel, C.C., et al., *Conserved principles of mammalian transcriptional regulation revealed by RNA half-life.* Nucleic Acids Res, 2009. **37**(17): p. e115.
46. Rabani, M., et al., *Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells.* Nat Biotechnol, 2011. **29**(5): p. 436-42.
47. Tani, H. and N. Akimitsu, *Genome-wide technology for determining RNA stability in mammalian cells: historical perspective and recent advantages based on modified nucleotide labeling.* RNA Biol, 2012. **9**(10): p. 1233-8.

48. Lindemann, S., et al., *Activated platelets mediate inflammatory signaling by regulated interleukin 1beta synthesis.* J Cell Biol, 2001. **154**(3): p. 485-90.
49. Zimmerman, G.A. and A.S. Weyrich, *Signal-dependent protein synthesis by activated platelets: new pathways to altered phenotype and function.* Arterioscler Thromb Vasc Biol, 2008. **28**(3): p. s17-24.
50. Cecchetti, L., et al., *Megakaryocytes differentially sort mRNAs for matrix metalloproteinases and their inhibitors into platelets: a mechanism for regulating synthetic events.* Blood, 2011. **118**(7): p. 1903-11.
51. Steiner, M. and M. Baldini, *Protein synthesis in aging blood platelets.* Blood, 1969. **33**(4): p. 628-33.
52. Gnatenko, D.V., et al., *Transcript profiling of human platelets using microarray and serial analysis of gene expression.* Blood, 2003. **101**(6): p. 2285-93.
53. McRedmond, J.P., et al., *Integration of proteomics and genomics in platelets: a profile of platelet proteins and platelet-specific genes.* Mol Cell Proteomics, 2004. **3**(2): p. 133-44.
54. Rondina, M.T., et al., *The septic milieu triggers expression of spliced tissue factor mRNA in human platelets.* J Thromb Haemost, 2011. **9**(4): p. 748-58.
55. Hawrylowicz, C.M., et al., *Activated platelets express IL-1 activity.* J Immunol, 1989. **143**(12): p. 4015-8.
56. Savini, I., et al., *Translational control of the ascorbic acid transporter SVCT2 in human platelets.* Free Radic Biol Med, 2007. **42**(5): p. 608-16.
57. Kieffer, N., et al., *Biosynthesis of major platelet proteins in human blood platelets.* Eur J Biochem, 1987. **164**(1): p. 189-95.
58. Weyrich, A.S., et al., *Signal-dependent translation of a regulatory protein, Bcl-3, in activated human platelets.* Proc Natl Acad Sci U S A, 1998. **95**(10): p. 5556-61.
59. Salzman, J., et al., *Cell-type specific features of circular RNA expression.* PLoS genetics, 2013. **9**: p. e1003777.
60. Bachmayr-Heyda, A., et al., *Correlation of circular RNA abundance with proliferation--exemplified with colorectal and ovarian cancer, idiopathic lung fibrosis, and normal human tissues.* Sci Rep, 2015. **5**: p. 8057.
61. Wu, C.S., et al., *Integrative transcriptome sequencing identifies trans-splicing events with important roles in human embryonic stem cell pluripotency.* Genome Res, 2014. **24**(1): p. 25-36.
62. Westholm, J.O., et al., *Genome-wide analysis of drosophila circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation.* Cell Rep, 2014. **9**(5): p. 1966-80.
63. Li, Y., et al., *Circular RNA is enriched and stable in exosomes: a promising biomarker for cancer diagnosis.* Cell Res, 2015. **25**(8): p. 981-4.

**LEGENDS TO TABLES AND FIGURES**

**Table 1. circRNA read support for structures and their frequency in all samples.** For each sample, the number of distinct circRNA structures supported by over 10, 100 and 1000 circRNA (back-splice) reads are shown, together with the number of circRNA reads per million reads (K/M). D – Digested with RNAseR, U – Undigested, matched sample.

**Figure 1. CircRNAs in human tissues**. The number of circRNAs (structure counts), junction reads (read counts), circRNA producing genes (gene counts), mean number of circRNA transcripts per circRNA producing genes, and RNAseq library sizes (read counts) are shown. Samples are presented as 4 groups: nucleated tissues, nucleated cell lines, nucleated cell lines digested with RNAseR, and anucleated cell types. The number of back spice reads is significantly higher in anucleate samples (platelets and RBCs) than all others ($p=2.7x10^4$ Wilcoxon Rank Sum test, figures adjusted for library size) All data was processed using PTESFinder using publicy available datasets, with the exception of the data from RBC (see Methods and Supplementary Tables S2 for details of datasets and S3 for details of structures).

**Figure 2. CircRNA structures per gene. A.** Histograms showing circRNA numbers per gene established using PTESFinder for RBCs (this study), Platelets [11], Fibroblasts and Fibroblasts digested with RNAseR [23]. **B.** Schematic diagram showing *XPO1* gene intron/exon organisation together with the read counts and inferred structure of all back-splice junctions identified within the 3 platelet samples. Inferred structures assume that all internal Refseq exons are present within the structure defined by each back-splice. Read counts for the 11 structures identified previously within RNAseR digested H9 ES cells [27] are highlighted in red.

**Figure 3. Confirmation of circRNA abundance and resistance to RNAseR**. **A.** Schema of qPCR assays using an E5-E2 circRNA as an example. All assays use a common reporter probe, and utilise either an exon downstream of the circRNA to assay linear expression (probe in donor exon), or an exon upstream (probe in acceptor exon). **B**. **Left panel** - Expression levels (-ΔCT values) of linear (Ex1-2) and circular (Ex5-2, and Ex4-2) MAN1A2 transcripts relative to housekeeping pool. **Right panel** - Expression levels of circRNAs relative to linear RNAs from the same loci normalised to housekeeping pool (-ΔΔCT values). **C.** Expression of 9 circRNAs relative to linear forms from the same loci, normalised to housekeeping pool (-ΔΔCT values). **D.** Change in CT values of circRNAs relative to linear forms from the same loci in RNAse digested RNAs, normalised to mock digested RNAs (-ΔΔCT values). Templates, circRNAs and linear forms assayed are indicated.

**Figure 4. Differential read depth defines genes with significant circRNA enrichment in platelets. A.** Box and Whisker plots showing the Ratio of RPKM from circRNA producing exons ($RPKM_I$) to RPKM from exons which do not produce circRNAs ($RPKM_E$) for all genes in each sample. The median, upper and lower quartiles are shown, with outliers as solid circles. **B.** The proportion of reads from circRNA producing exons averaged across all nucleated samples (Y axis) and platelets (X axis). **C.** Fold enrichment of reads from circRNA producing exons in platelets relative to nucleated tissues. All genes with an average $RPKM_I$ >1 in platelets and expressed in 8 or more nucleated tissues are shown. Blue – genes significantly enriched in platelets. Red – genes not significantly enriched in platelets. The datapoints corresponding to the 5 most enriched genes are indicated. The slope x=y is shown as a dashed line.

**Figure 5. circRNA enrichment occurs in platelets but not megakaryocytes. A-D.** adapted UCSC screen shots for 4 genes which extreme circRNA enrichment in platelets. Each panel shows RNAseq read abundance in a single Platelet total RNA sample (F1), together with the position and abundance of back-splice exon junctions in the same sample. The intron-exon structure is shown below, with exons known to contribute to circRNAs shown in red. CircRNA structures present in the sample together with back-splice frequencies are also shown (for *PHC3*, only structures with 10 or more

back-splice reads are shown). Above each panel, the gene name, cytogenetic location, coding strand, and scale in kilobases, is indicated. **E-F.** Expression levels of circRNAs relative to linear RNAs from the same loci in platelets and cultured megakaryocytes, normalised to housekeeping pool (-ΔΔCT values).

**Figure 6. Degradation of platelet RNA**. **A.** Correlation of the change in gene expression between megakaryocte polyA+ RNA and platelet polyA+ RNA (Y axis), and mRNA half life estimates (X axis). Distributions of both are shown as histograms on each axis. **B.** qPCR analysis of differential decay of linear and circRNAs in platelets following incubation at 37$^{o}$c for 0, 72 and 96 hours. Data from 3 biological replicates are shown. **Top left panel** - Expression levels of housekeeping genes (CT values); **Top right panel** - Expression levels of linear structures from 5 circRNA producing genes relative to housekeeping pool (-ΔCT values); **Bottom panel** - Expression levels of 7 circRNAs relative to linear transcripts from the same loci, both normalised to housekeeping pool (-ΔΔCT values).

**Table 1**

| circRNA reads | Bladder | Liver | Skin | Muscle | Kidney | Breast | Prostate | Brain | Heart | Ovary | Lung | Colon | K562_1 | GM12878_1 | K562_2 | GM12878_2 | Fibroblasts_U1 | Fibroblasts_U2 | H9_D | Fibroblasts_D1 | Fibroblasts_D2 | Platelet_F | Platelet_M2 | Platelet_M1 | Red Blood Cells |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **>10** | 55 | 308 | 164 | 136 | 196 | 186 | 71 | 670 | 235 | 263 | 521 | 190 | 3 | 9 | 11 | 6 | 125 | 141 | 227 | 2797 | 3052 | 7281 | 9605 | 8459 | 1272 |
| **>100** | 0 | 11 | 3 | 2 | 2 | 3 | 1 | 19 | 5 | 3 | 6 | 5 | 0 | 0 | 0 | 0 | 4 | 2 | 2 | 200 | 224 | 1197 | 1686 | 1325 | 96 |
| **>1000** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 2 | 5 | 83 | 147 | 95 | 2 |
| **K/M** | 66 | 295 | 170 | 69 | 104 | 126 | 59 | 447 | 148 | 125 | 172 | 84 | 9 | 7 | 9 | 9 | 45 | 36 | 427 | 546 | 464 | 8469 | 11138 | 7579 | 597 |

Figure 1

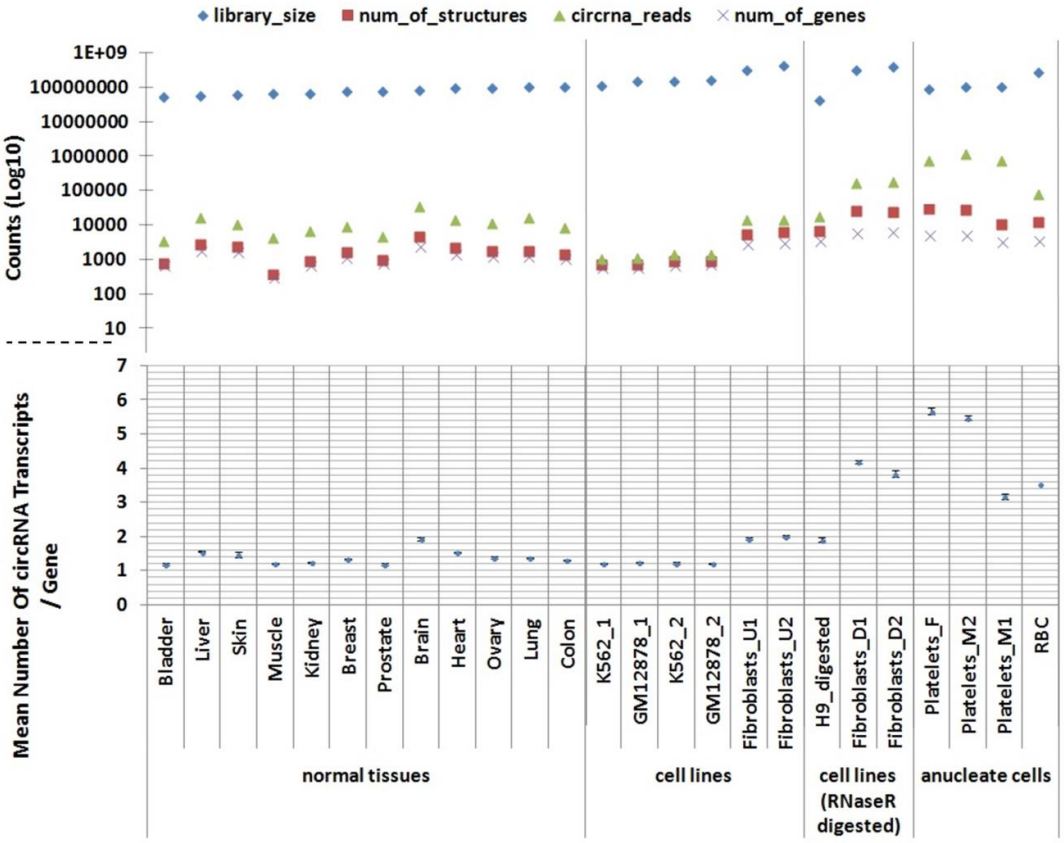


Figure 2

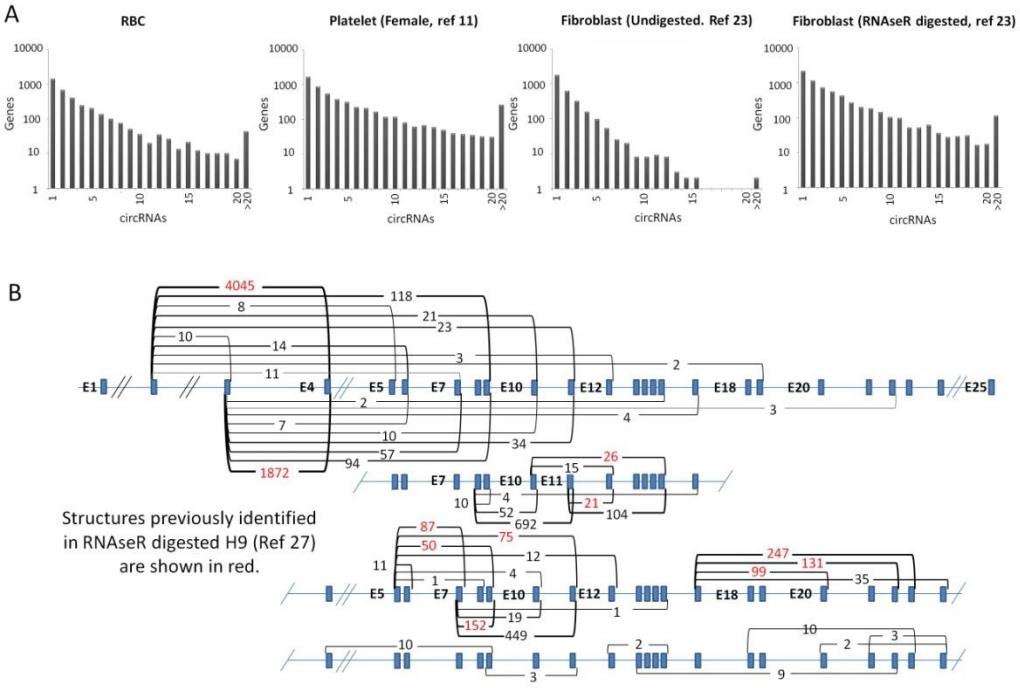
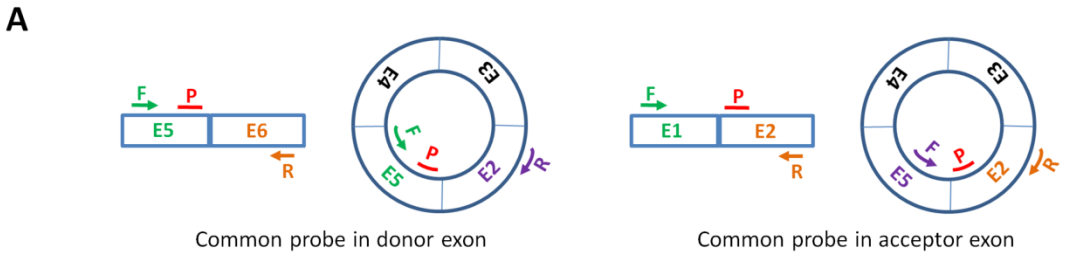
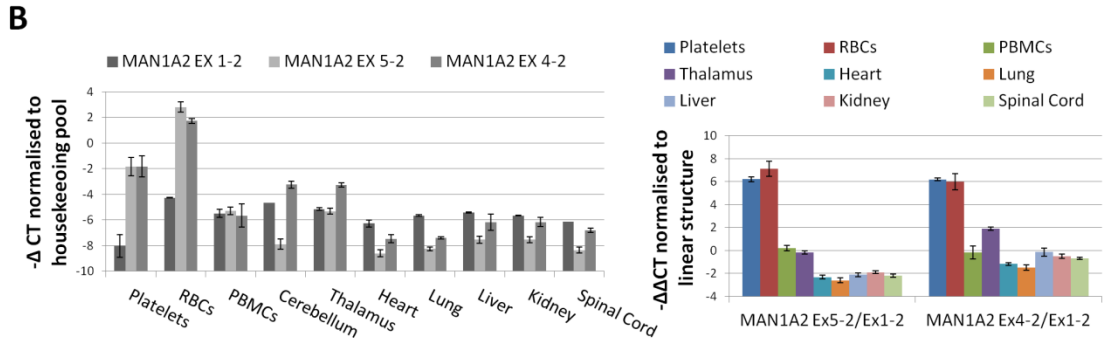
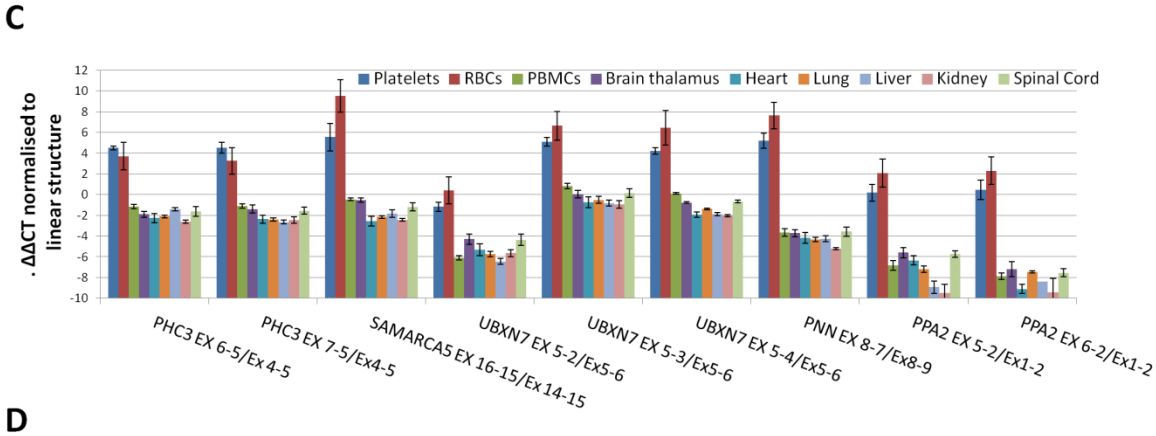

A
RBC
Platelet (Female, ref 11)
Fibroblast (Undigested. Ref 23)
Fibroblast (RNAseR digested, ref 23)
Genes
circRNAs
B
E1
E4
E5
E7
E10
E11
E12
E18
E20
E25
Structures previously identified in RNAseR digested H9 (Ref 27) are shown in red.

Figure 3


A
F
P
E5
E6
R
E4
E3
E2
E5
Common probe in donor exon
F
P
E1
E2
R
E4
E3
E2
E5
Common probe in acceptor exon


B
MAN1A2 EX 1-2
MAN1A2 EX 5-2
MAN1A2 EX 4-2
-Δ CT normalised to housekeeoing pool
Platelets
RBCs
PBMCs
Cerebellum
Thalamus
Heart
Lung
Liver
Kidney
Spinal Cord
Platelets
RBCs
PBMCs
Thalamus
Heart
Lung
Liver
Kidney
Spinal Cord
-ΔΔCT normalised to linear structure
MAN1A2 Ex5-2/Ex1-2
MAN1A2 Ex4-2/Ex1-2


C
. ΔΔCT normalised to linear structure
Platelets
RBCs
PBMCs
Brain thalamus
Heart
Lung
Liver
Kidney
Spinal Cord
PHC3 EX 6-5/Ex 4-5
PHC3 EX 7-5/Ex4-5
SAMARCA5 EX 16-15/Ex 14-15
UBXN7 EX 5-2/Ex5-6
UBXN7 EX 5-3/Ex5-6
UBXN7 EX 5-4/Ex5-6
PNN EX 8-7/Ex8-9
PPA2 EX 5-2/Ex1-2
PPA2 EX 6-2/Ex1-2

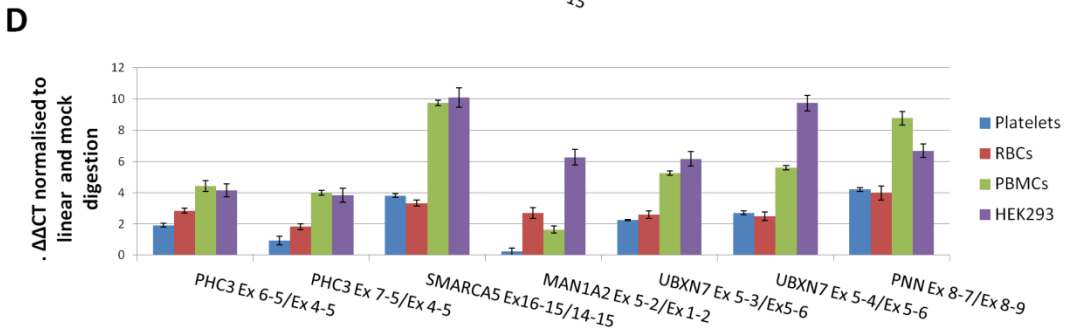
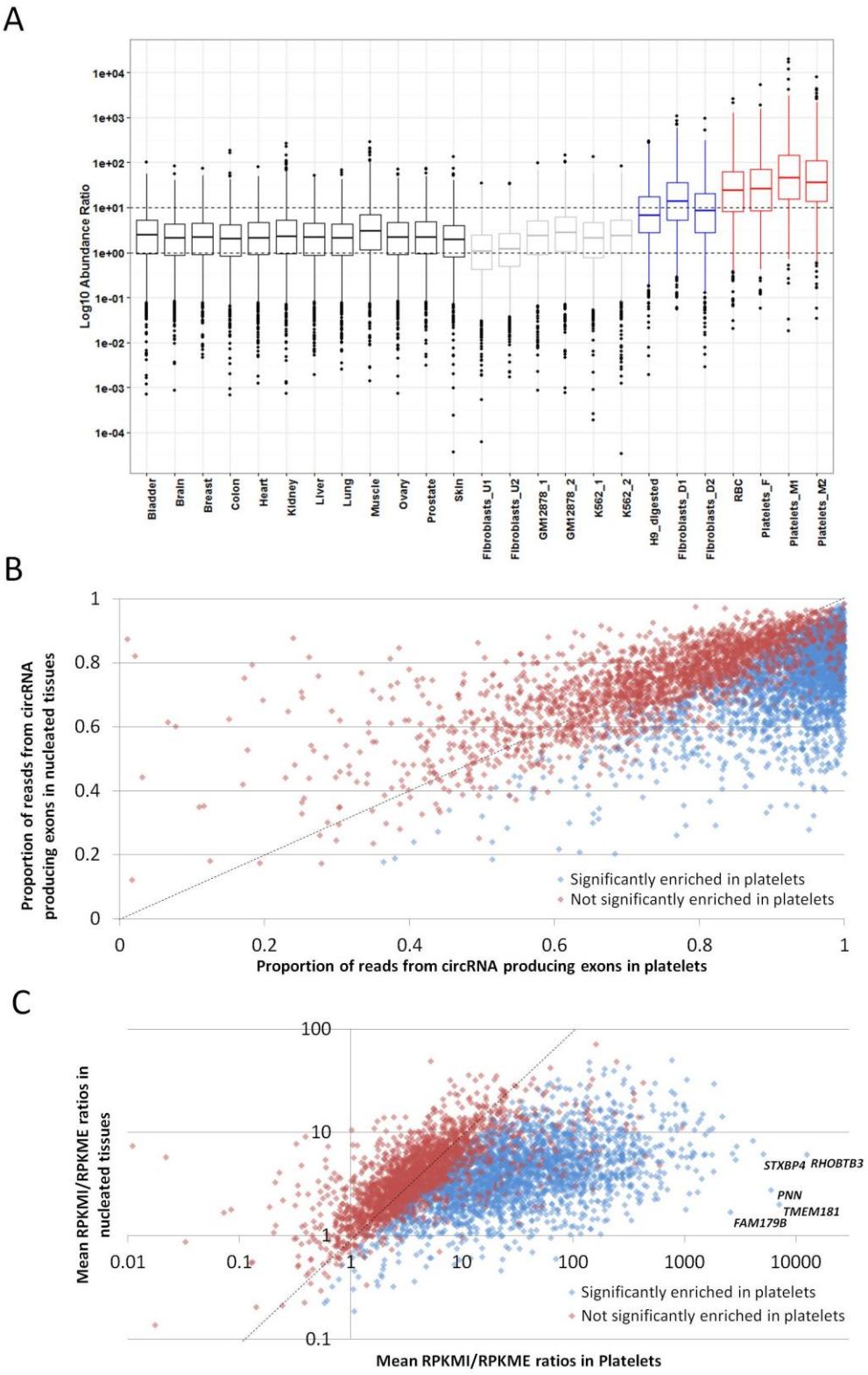

D
. ΔΔCT normalised to linear and mock digestion
Platelets
RBCs
PBMCs
HEK293
PHC3 Ex 6-5/Ex 4-5
PHC3 Ex 7-5/Ex 4-5
SMARCA5 Ex16-15/14-15
MAN1A2 Ex 5-2/Ex 1-2
UBXN7 Ex 5-3/Ex5-6
UBXN7 Ex 5-4/Ex 5-6
PNN Ex 8-7/Ex 8-9

**Figure 4. Enrichment of reads from circRNA producing exons**



# Figure 5

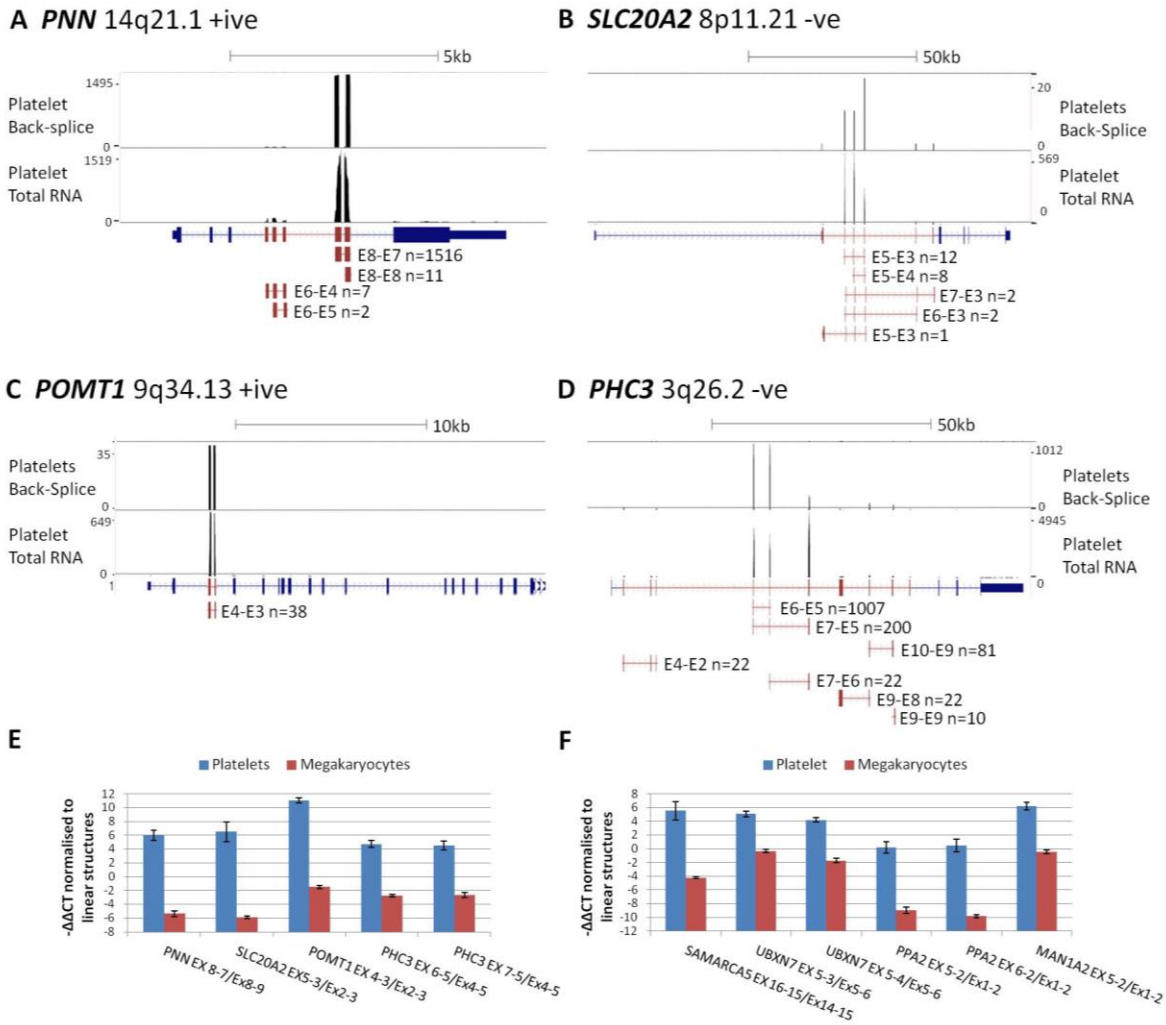
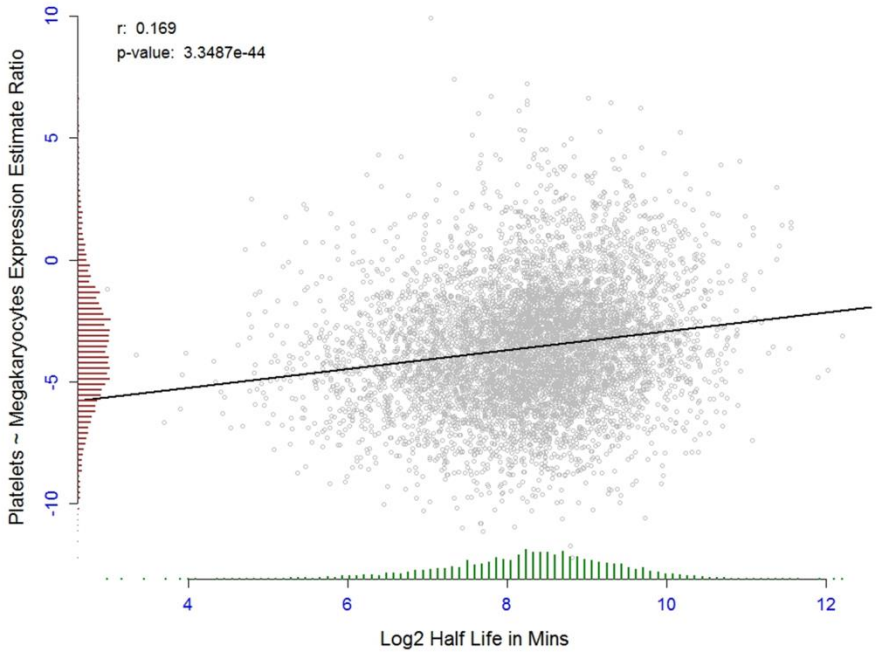
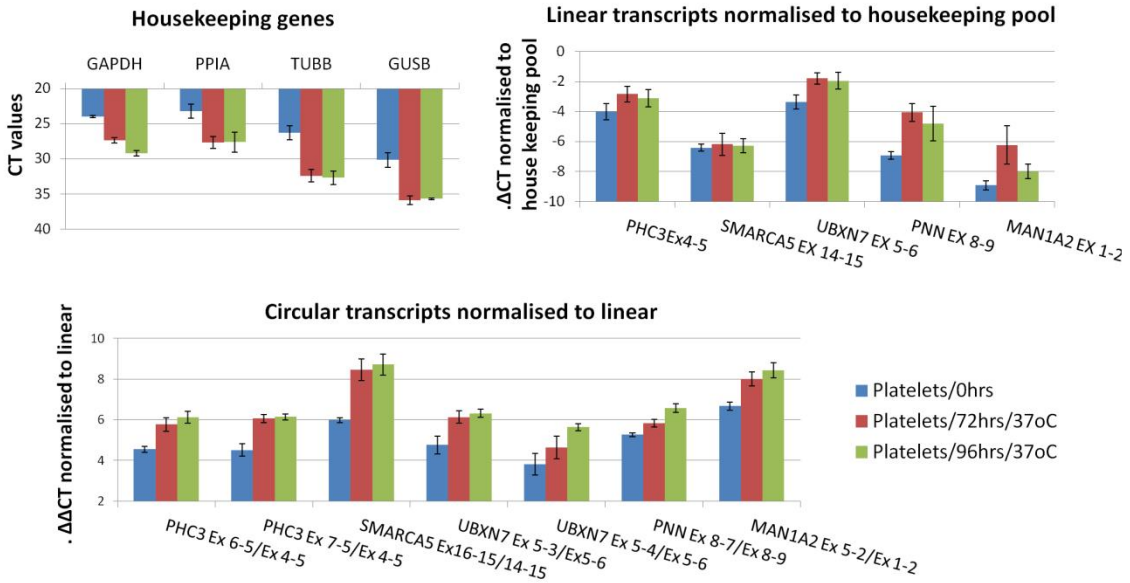

A PNN 14q21.1 +ive
5kb
1495
Platelet
Back-splice
0
1519
Platelet
Total RNA
0
E8-E7 n=1516
E8-E8 n=11
E6-E4 n=7
E6-E5 n=2
B SLC20A2 8p11.21 -ve
50kb
20
Platelets
Back-Splice
0
569
Platelet
Total RNA
0
E5-E3 n=12
E5-E4 n=8
E7-E3 n=2
E6-E3 n=2
E5-E3 n=1
C POMT1 9q34.13 +ive
10kb
35
Platelets
Back-Splice
0
649
Platelet
Total RNA
0
1
E4-E3 n=38
D PHC3 3q26.2 -ve
50kb
1012
Platelets
Back-Splice
0
4945
Platelet
Total RNA
0
E6-E5 n=1007
E7-E5 n=200
E10-E9 n=81
E4-E2 n=22
E7-E6 n=22
E9-E8 n=22
E9-E9 n=10
E
Platelets
Megakaryocytes
-ΔΔCT normalised to linear structures
12
10
8
6
4
2
0
-2
-4
-6
-8
PNN EX 8-7/Ex8-9
SLC20A2 EX5-3/Ex2-3
POMT1 EX 4-3/Ex2-3
PHC3 EX 6-5/Ex4-5
PHC3 EX 7-5/Ex4-5
F
Platelet
Megakaryocytes
-ΔΔCT normalised to linear structures
8
6
4
2
0
-2
-4
-6
-8
-10
-12
SAMARCA5 EX 16-15/Ex14-15
UBXN7 EX 5-3/Ex5-6
UBXN7 EX 5-4/Ex5-6
PPA2 EX 5-2/Ex1-2
PPA2 EX 6-2/Ex1-2
MAN1A2 EX 5-2/Ex1-2

Figure 6

**A**


r: 0.169
p-value: 3.3487e-44
Platelets ~ Megakaryocytes Expression Estimate Ratio
10
5
0
-5
-10
4
6
8
10
12
Log2 Half Life in Mins

**B**


Housekeeping genes
GAPDH
PPIA
TUBB
GUSB
CT values
20
25
30
35
40
Linear transcripts normalised to housekeeping pool
.ΔCT normalised to house keeping pool
0
-2
-4
-6
-8
-10
PHC3Ex4-5
SMARCA5 EX 14-15
UBXN7 EX 5-6
PNN EX 8-9
MAN1A2 EX 1-2
Circular transcripts normalised to linear
. ΔΔCT normalised to linear
10
8
6
4
2
PHC3 Ex 6-5/Ex 4-5
PHC3 Ex 7-5/Ex 4-5
SMARCA5 Ex16-15/14-15
UBXN7 Ex 5-3/Ex5-6
UBXN7 Ex 5-4/Ex 5-6
PNN Ex 8-7/Ex 8-9
MAN1A2 Ex 5-2/Ex 1-2
Platelets/0hrs
Platelets/72hrs/37oC
Platelets/96hrs/37oC